AN OPTIMAL PRIOR KNOWLEDGE-BASED DOA ESTIMATION METHOD

Guillaume Bouleux [1], *Petre Stoica* [2] *and Rémy Boyer* [3]

[1] LASPI
Saint Etienne University
IUT de Roanne, Roanne, France
guillaume.bouleux@univ-st-etienne.fr

[2] Systems and Control Division
Uppsala University
Uppsala, Sweden
peter.stoica@it.uu.se

[3] LSS-Supelec
Paris-Sud University
Gif-Sur-Yvette, France
remy.boyer@lss.supelec.fr

ABSTRACT

In some applications involving line spectra or direction of arrival (DOA) estimation, we may have a priori information which could consist of known frequencies in Magnetoencephalography (MEG) or mechanical signals, or of known DOA's in a RADAR urban scenario. With this fact in mind we propose an optimal method for array processing that exploits the information of the known DOA's for estimating the unknown DOA's as accurately as possible. This new **P**rior-know**LEDGE** (**PLEDGE**) technique is based on the method of direction estimation (MODE) approach. To show the benefits of incorporating prior-knowledge we also present the corresponding stochastic Cramér-Rao Bound (CRB). Finally, we use PLEDGE for estimating frequencies in the current spectrum of an induction motor to perform the diagnosis on rotor bars.

1. INTRODUCTION

In some applications involving direction of arrival (DOA) or line spectra estimation, we may have known a priori information. For example, in a RADAR urban scenario, the presence of stationary sources such as buildings, leads to waves that impinge on the array from known directions. In such a multi-path problem, the sources are highly correlated or even coherent. Similarly, in Magnetoencephalography (MEG) some neural activities of the brain are repetitive and thus we know frequencies present in the synaptic current spectrum. A more industrial application but still in the same context is the diagnosis of rotating machines. The kinematic of these systems is well known and particular frequencies are well identified, which could be on the line signal frequency or gear frequency. The point is that these prior knowledge might not carry any information and might mask what is of interest. It is then crucial to take advantage of the prior-information for estimating the parameters of interest as accurately as possible. Following this way, recent work has been proposed in the context of NMR spectroscopy [4] and array processing [8]. Each of these methods uses the concept of reduced signal subspace by deflating the sample covariance matrix. The Cramér-Rao Bound (CRB) associated with this model has been given and studied in [2]. This bound, the Prior-CRB, shows that this kind of methods is limited and suffers from drastic assumptions as well. We can argue that the non-bijective transformation (the deflation) causes an information loss. From this it is obvious and easy to understand why these algorithms are limited.

The idea pursued in this paper is to propose an optimal algorithm in which we could easily use and fix the known parameters. We have based our derivation on the popular method of direction estimation (MODE) [12] from which we have deduced the optimal **P**rior-know**LEDGE** (**PLEDGE**) algorithm for array processing. The PLEDGE is presented for the Uniform and Linear Array (ULA) case which allows us to easily transpose the algorithm to the line spectra estimation. To show the benefits of incorporating prior-knowledge we also present the corresponding Cramér-Rao Bound (CRB). Based on simulation results we present a discussion on what prior-hypothesis is essential. We end this work by presenting a diagnosis problem on rotor bars in an induction motor.

2. MODEL STATEMENT

Consider n narrowband and far-field plane waves impinging on an ULA composed of L sensors separated by a half wavelength. Let t be a sample ("snapshot") and assume that the total number of available samples is N, then $t = 1, \dots N$. The one sample response, or equivalently the single-experiment time series model can be parametrized in the following way

$$\boldsymbol{y}(t) = \boldsymbol{A}(\boldsymbol{\omega})\boldsymbol{x}(t) + \boldsymbol{n}(t) \tag{1}$$

where $\boldsymbol{A}(\boldsymbol{\omega}) = [\boldsymbol{a}(\omega_1) \dots \boldsymbol{a}(\omega_n)]$ and where $\boldsymbol{a}(\omega_i)$ is the i-th steering vector defined by

$$\boldsymbol{a}(\omega_i) = [1 \; e^{j\omega_i} \dots e^{j(L-1)\omega_i}]^T \tag{2}$$

with $\omega_i = -\pi \sin(\theta_i)$ the spatial pulsation for the direction of arrival (DOA) estimation problem (θ the DOA) and where ω_i could be the temporal pulsation, i.e $\omega_i = 2\pi \frac{f_i}{f_e}$ (f_e the sample frequency) for the line spectra estimation problem. Amplitude waves $\boldsymbol{x}(t)$ and noise signal $\boldsymbol{n}(t)$ are assumed to be jointly Gaussian with zero-mean, stationary and circular stochastic process of second order moments

$$\mathbb{E}[\boldsymbol{x}(t)\boldsymbol{x}^H(t)] = \boldsymbol{P} \quad \textit{and} \quad \mathbb{E}[\boldsymbol{n}(t)\boldsymbol{n}^H(t)] = \sigma^2\boldsymbol{I} \tag{3}$$

where $\mathbb{E}$ stands for the mathematical expectation and ^{H} for the transpose and conjugate. For the latter use we introduce ^{T} as the transpose and * as the conjugate.

The aim of this work is to furnish an optimal method for the DOA estimation which integrates the prior-knownledge of some directions. The optimality has to be understood in the Maximum Likelihood (ML) sense and elaborating ML DOA estimation has been treated in abundance since the three last decades (see for instance [9]). MODE rose up particularly by its good performances and computational cost (among other things) [10, 6, 12]. We briefly recall its principle before using it to tackle the problem of prior-information integration.

3. MODE

The derivation of MODE can be seen by different points of view and we choose a pragmatic manner without going into depth, for more details you can refer to [12] and the references therein. Assume that the directions can be found by polynomial rooting. Consequently, we define the polynomial

$$b_0 z^n + b_1 z^{n-1} + \dots + b_n = b_0 \prod_{i=1}^{n} (z - e^{j\omega_i}) \tag{4}$$

in which the n angle roots are the n directions of interest. Since all the roots of (4) belong to the unit circle, its coefficients must satisfy the conjugate symmetry constraint [12], and thus $b_i = b^*_{n-i}$, $i = 0, \dots, n$. Let us define now the Sylvester matrix whose rows are formed by the polynomial coefficients $\{b_i\}$ such as

$$\boldsymbol{B}^H = \begin{bmatrix} b_n & b_{n-1} & \dots & b_0 & & \mathbf{0} \\ & \ddots & \ddots & & \ddots & \\ \mathbf{0} & & b_n & b_{n-1} & \dots & b_0 \end{bmatrix} \tag{5}$$

meaning that the columns of $\boldsymbol{B}$ span the null space of $\boldsymbol{A}$.
Consider now by deterministic derivation, the concentrated negative log-likelihood function [9] $F_{ML}(\boldsymbol{\omega}) = Trace[\boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp}\hat{\boldsymbol{R}}]$ or equivalently

$$F_{ML}(\boldsymbol{b}) = Trace[\boldsymbol{B}(\boldsymbol{B^H B})^{-1}\boldsymbol{B^H}\hat{\boldsymbol{R}}] \tag{6}$$

with $\boldsymbol{b} = [b_0 \ \dots \ b_n]^T$, where $\boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp} = \boldsymbol{I} - \boldsymbol{A}\left(\boldsymbol{A^H A}\right)^{-1}\boldsymbol{A^H}$, and where the sample covariance matrix is defined by

$$\hat{\boldsymbol{R}} = \frac{1}{N}\sum_{t=1}^{N}\boldsymbol{y}(t)\boldsymbol{y}^H(t). \tag{7}$$

Let

$$\hat{\boldsymbol{R}} = \hat{\boldsymbol{E}}_S\hat{\boldsymbol{\Sigma}}_S\hat{\boldsymbol{E}}_S^H + \hat{\boldsymbol{E}}_N\hat{\boldsymbol{\Sigma}}_N\hat{\boldsymbol{E}}_N^H \tag{8}$$

be an eigendecomposition of $\hat{\boldsymbol{R}}$. Here, $\hat{\boldsymbol{\Sigma}}_S$ is a diagonal matrix containing the n largest eigenvalues, and the columns of $\hat{\boldsymbol{E}}_S$ are the corresponding eigenvectors. Similarly, $\hat{\boldsymbol{\Sigma}}_N$ contains the $L-n$ remaining eigenvalues and the columns of $\hat{\boldsymbol{E}}_N$ are the associated eigenvectors. A consistent estimate of the noise variance is then given by $\hat{\sigma}^2 = \frac{1}{L-n}Trace[\hat{\boldsymbol{\Sigma}}_N]$.
It has been shown that the ML estimate was computationally heavy and not efficient, i.e. the variance of the estimation errors $(\hat{\omega}_i - \omega_i)$ do not achieve the CRB for a small number of sensor [12]. So, a better approach is given by a large sample realization of (6) given by

$$Trace[\boldsymbol{A}^H\hat{\boldsymbol{E}}_N\hat{\boldsymbol{E}}_N^H\boldsymbol{A}\boldsymbol{W}] \tag{9}$$

where $\boldsymbol{W}$ is a positive definite weighting matrix. The derivation of the optimal (ML sense) $\boldsymbol{W}$ can be found in [12] and after algebraic manipulations we obtain the MODE minimizer. The coefficients $\{b_i\}$ are therefore the solution of the minimizer [12]

$$F_{Mode}(\boldsymbol{b}) = Trace[\ \boldsymbol{B}(\hat{\boldsymbol{B}}^H\hat{\boldsymbol{B}})^{-1}\boldsymbol{B}^H\hat{\boldsymbol{E}}_S\hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{E}}_S^H] \tag{10}$$

with

$$\hat{\boldsymbol{\Lambda}} = \hat{\boldsymbol{\Sigma}}_S(\hat{\boldsymbol{\Sigma}}_S - \hat{\sigma}^2\boldsymbol{I})^{-2}. \tag{11}$$

We make use of the rank-1 equivalence of $\hat{\boldsymbol{E}}_S\hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{E}}_S^H$ to observe that (10) can be viewed as a quadratic minimizer. To this end, we use conjointly

$$\hat{\boldsymbol{E}}_S\hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{E}}_S^H = \sum_{k=1}^{n}\hat{\lambda}_k\hat{\boldsymbol{e}}_k\hat{\boldsymbol{e}}_k^H \tag{12}$$

with $\hat{\boldsymbol{e}}_k$ the k-th column of $\hat{\boldsymbol{E}}_S$ and its corresponding eigenvalue $\hat{\lambda}_k$, and property

$$\boldsymbol{B}^H\hat{\boldsymbol{e}}_k = \hat{\tilde{\boldsymbol{E}}}_k\boldsymbol{b} \tag{13}$$

introduced in [3] and where $\hat{\tilde{\boldsymbol{E}}}_k$ is a $(L-n+1)\times(n+1)$ Hankel matrix defined such as

$$\hat{\tilde{\boldsymbol{E}}}_k = \begin{bmatrix} \hat{e}_k(0) & \hat{e}_k(1) & \dots & \hat{e}_k(n) \\ \hat{e}_k(1) & \dots & \dots & \hat{e}_k(n+1) \\ \vdots & & & \vdots \\ \hat{e}_k(L-n) & \dots & \dots & \hat{e}_k(L) \end{bmatrix}, \tag{14}$$

to remark that (10) can effectively be written

$$Trace[\ \boldsymbol{B}(\hat{\boldsymbol{B}}^H\hat{\boldsymbol{B}})^{-1}\boldsymbol{B}^H\hat{\boldsymbol{E}}_S\hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{E}}_S^H] = \boldsymbol{b}^H\hat{\boldsymbol{Q}}\boldsymbol{b} \tag{15}$$

where

$$\hat{\boldsymbol{Q}} = \sum_{k}^{n}\hat{\lambda}_k\ \hat{\tilde{\boldsymbol{E}}}_k^H(\hat{\boldsymbol{B}}^H\hat{\boldsymbol{B}})^{-1}\hat{\tilde{\boldsymbol{E}}}_k. \tag{16}$$

As we can see from (16), we have to first give an estimate of $\hat{\boldsymbol{B}}$. It has been proved in [12] that an initial consistent estimate $\hat{\boldsymbol{b}}$ is obtained by taking $(\hat{\boldsymbol{B}}^H\hat{\boldsymbol{B}}) = \boldsymbol{I}$. Constraining the minimizer with $\|\boldsymbol{b}\|^2 = 1$, we obtained $\hat{\boldsymbol{b}}$ as the eigenvector corresponding to the smallest eigenvalue of $\hat{\boldsymbol{Q}} = \sum_k^n\hat{\lambda}_k\ \hat{\tilde{\boldsymbol{E}}}_k^H\hat{\tilde{\boldsymbol{E}}}_k$. Making use of this consistent estimate, we form $\hat{\boldsymbol{B}}$ and (15) is solved with the constraint $\|\boldsymbol{b}\|^2 = 1$ by selecting the eigenvector associated with the smallest eigenvalue of (16). Lastly, the n DOA's are deduced from the n angle roots of (4).

4. PLEDGE SCHEME

We deduce PLEDGE directly from MODE. Thanks to the polynomial rooting, we can easily write a new polynomial which integrates known zeros. In such a case, the polynomial defined in (4) can be factorized in the following manner

$$b_0\prod_{i=1}^{n}(z - e^{j\omega_i}) = \mathcal{Q}_k(z)\mathcal{Q}_u(z) \tag{17}$$

where the two following polynomials are introduced

$$\mathcal{Q}_k(z) \quad = \quad q_0\prod_{i=1}^{n_k}(z - e^{j\omega_i}) = q_0 z^{n_k} + \dots + q_{n_k} \tag{18}$$

$$\mathcal{Q}_u(z) \quad = \quad \bar{b}_0\prod_{i=1}^{n_u}(z - e^{j\omega_i}) = \bar{b}_0 z^{n_u} + \dots + \bar{b}_{n_u}, \tag{19}$$

which are respectively the polynomials whose zeros are the known and the unknown DOA's and where without loss of generality [12], the modulus of both q_0 and $\bar{b}_0$ are arbitrary. The $\boldsymbol{b}$ vector can therefore be formulated in algebraic correspondence. It is enough for this to convolve the coefficients of the two polynomials $\mathcal{Q}_k(z)$ and $\mathcal{Q}_u(z)$ and it yields

$$[b_0 \quad b_1 \quad \dots \quad b_n]^T = \boldsymbol{C}^T\ \bar{\boldsymbol{b}} \tag{20}$$

where

$$\begin{aligned} \bar{\boldsymbol{b}} &= \begin{bmatrix}\bar{b}_0 & \bar{b}_1 & \dots\bar{b}_{n_u}\end{bmatrix}^T \\ \boldsymbol{C} &= \begin{bmatrix} q_0 & q_1 & \dots & q_{n_k} & & \mathbf{0} \\ & \ddots & \ddots & & \ddots & \\ \mathbf{0} & & q_0 & q_1 & \dots & q_{n_k} \end{bmatrix}. \end{aligned}$$

PLEDGE estimates the n_u unknown DOA's from the estimated polynomial (19). To solve this problem, we substitute (20) into (15) and the following modified quadratic optimization problem is deduced

$$Trace[\boldsymbol{B}^H(\hat{\boldsymbol{B}}^H\hat{\boldsymbol{B}})^{-1}\hat{\boldsymbol{E}}_S\hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{E}}_S^H\boldsymbol{B}] = \bar{\boldsymbol{b}}^H\hat{\bar{\boldsymbol{Q}}}\bar{\boldsymbol{b}} \tag{21}$$

with $\hat{\bar{\boldsymbol{Q}}} = \boldsymbol{C}\hat{\boldsymbol{Q}}\boldsymbol{C}^H$. From here we follow the steps required for finding the DOA's from MODE, namely :

1) find a first consistent estimate of (21) by taking $(\hat{\boldsymbol{B}}^H\hat{\boldsymbol{B}}) = \boldsymbol{I}$ and constrain the minimizer with $\|\boldsymbol{b}\|^2 = 1$. We stress on the fact that as it might not be an evidence, the best performances are obtained when the constraint is equal to $\|\boldsymbol{b}\|^2 = 1$ and not $\|\bar{\boldsymbol{b}}\|^2 = 1$. We will not go further on this here,
2) form the $\hat{\boldsymbol{B}}$ matrix and solve (21) subject to $\|\boldsymbol{b}\|^2 = 1$, then finally
3) the n_u DOA's are determined as the angle of the roots of (19) whose coefficients are the $\{\hat{\hat{b}}_i\}$ estimated at step 2).

5. PLEDGE CRB

The CRB proposed in [2] is based on a subspace reduction by an orthogonal projection method which guarantees to suppress the known information. This bound, named the Prior-CRB (P-CRB), has been derived under the deterministic assumption. The results and analysis which were given, pointed out that exploiting the prior-information by an orthogonal projection was not an optimal method. Actually the main observations are : (1) the $\textit{P-CRB} = \textit{CRB}$ in general conditions and (2) the $\textit{P-CRB} < \textit{CRB}$ when sources associated with the known and unknown directions are correlated. So, in such a procedure, the only way to improve the estimation of the unknown DOA's is to be in presence of full correlated sources. We can advance two reasons for that. One is surely the determinist derivation of the P-CRB, and the other could be felt by the fact we project the data leading to an irremediable information loss. Starting from this point, we propose in this section the stochastic CRB based on PLEDGE model and thereby named PLEDGE CRB.

We drew our inspiration from [11] to formulate and derive the PLEDGE CRB. Accordingly, let

$$\boldsymbol{\alpha} = [\boldsymbol{\omega}_u^T \; \boldsymbol{\rho}^T \; \sigma^2]^T \tag{22}$$

be the unknown parameter vector, where $\boldsymbol{\omega}_u = [\omega_1 \; \ldots \; \omega_u]^T$. Actually $\boldsymbol{\omega}_u$ is the set of unknown DOA's and indices here have no importance, they are arbitrary and help the mathematical formulation. The $n^2 \times 1$ vector $\boldsymbol{\rho}$ is made from $\{\boldsymbol{P}_{ii}\}$ and $\{Re(\boldsymbol{P}_{ij}), Im(\boldsymbol{P}_{ij}) \text{ for } j > i\}$. Under the previous assumptions and the Gaussian hypothesis, the Fisher Information Matrix (FIM) for the parameter vector $\boldsymbol{\alpha}$ is given by

$$FIM_{p,q} = N\,Trace\left[\boldsymbol{R}^{-1}\frac{\partial \boldsymbol{R}}{\partial \alpha_p}\boldsymbol{R}^{-1}\frac{\partial \boldsymbol{R}}{\partial \alpha_q}\right] \tag{23}$$
$$\text{for } p,q = 1,\ldots,n_u + n^2 + 1$$

with the true covariance matrix $\boldsymbol{R} = \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^H + \sigma^2\boldsymbol{I}$. The further derivation steps require first the vectorization of (23)

$$\frac{1}{N}FIM = \left(\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\alpha}^T}\right)^H (\boldsymbol{R}^{-T} \otimes \boldsymbol{R}^{-1})\left(\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\alpha}^T}\right) \tag{24}$$

with $\otimes$ the kronecker product and

$$\boldsymbol{r} = vec(\boldsymbol{R}) = (\boldsymbol{A}^* \otimes \boldsymbol{A})vec(\boldsymbol{P}) + \sigma^2 vec(\boldsymbol{I}), \tag{25}$$

and second the following partitioning

$$(\boldsymbol{R}^{-T/2} \otimes \boldsymbol{R}^{-1/2})\left[\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\omega}_u^T} | \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\rho}^T}\,\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\sigma}^2}\right] = [\boldsymbol{G}_u|\boldsymbol{\Delta}] \tag{26}$$

which leads to

$$\begin{aligned}\frac{1}{N}FIM &= \begin{bmatrix}\boldsymbol{G}_u^H \\ \boldsymbol{\Delta}^H\end{bmatrix}[\boldsymbol{G}_u\boldsymbol{\Delta}] \\ &= \begin{bmatrix}\boldsymbol{G}_u^H\boldsymbol{G}_u & \boldsymbol{G}_u^H\boldsymbol{\Delta} \\ \boldsymbol{\Delta}^H\boldsymbol{G}_u & \boldsymbol{\Delta}^H\boldsymbol{\Delta}\end{bmatrix}.\end{aligned} \tag{27}$$

We conduct the derivation for the upper left submatrix $\boldsymbol{G}_u^H\boldsymbol{G}_u$ along with the methodology of [11]. With no surprise, the final result, agreeing with an intuitive approach, gives us an expression of the PLEDGE CRB

$$\begin{aligned}&\textit{PLEDGE CRB}(\boldsymbol{\omega}_u) = \\ &\frac{\sigma^2}{2N}\left[Re\left\{\left(\boldsymbol{D}_u^H\boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp}\boldsymbol{D}_u\right) \odot \left(\boldsymbol{P}_u^H\boldsymbol{A}^H\boldsymbol{R}^{-1}\boldsymbol{A}\boldsymbol{P}_u\right)^T\right\}\right]^{-1}\end{aligned} \tag{28}$$

with $\odot$ the Hadamard product and where the partitioning $\boldsymbol{P} = [\boldsymbol{P}_u|\boldsymbol{P}_k]$ is considered. Once again this is an arbitrary choice and the $\boldsymbol{P}_u$ matrix is composed by the columns of $\boldsymbol{P}$ associated with the sources whose directions are unknown. Lastly, $\boldsymbol{D}_u = [d_1 \ldots d_{n_u}]$ where $d_i = (d\boldsymbol{a}(\omega_i)/d\omega_i)$ is the first derivative of $\boldsymbol{a}(\omega)$ considered at ω_i.

6. SIMULATIONS

In all the scenarios considered, we have two sources impinging on a ULA with a half-wavelength distance between the sensors. The sources have the directions θ_1 and θ_2 respectively. We define the power of each source with respect to the correlation matrix as

$$\frac{\boldsymbol{P}}{\sigma^2} = \begin{bmatrix} 10^{\frac{SNR_1}{10}} & \rho \\ \rho^* & 10^{\frac{SNR_2}{10}}\end{bmatrix}$$

where σ^2 is kept equal to one for each simulation and ρ is the correlation coefficient. The DOA of interest is the second one, namely θ_2. The number of sensors, snapshots, the power of the sources are varying along with the figures but the two DOA's are located at $\theta_1 = 10°$ and $\theta_2 = 12°$ for all experiments concerning Fig.1. Lastly each results are the mean of 1000 independent trials. We compare PLEDGE to MODE and P-MUSIC [2] except we have improved the performance of the latter by using the noise-free sample covariance matrix instead of directly the sample one (i.e. $\hat{\boldsymbol{R}} - \hat{\sigma}^2\boldsymbol{I}$ instead of $\hat{\boldsymbol{R}}$). We have plotted the stochastic CRB for n parameters, the PLEDGE CRB, the CRB for uncorrelated sources of Jansson et *al* [7] and the CRB for n_u unknown DOA's, $n_u = 1$ in the context of these simulations.

6.1. On PLEDGE performances

First, PLEDGE achieves the PLEDGE CRB whatever be the scenario. This fact concludes that PLEDGE is an optimal solution to the prior knowledge-based DOA estimation. Next, the set of Fig.1 and Fig.2's plots show the benefit of using PLEDGE instead of MODE for low and moderate SNR. For example, one would wonder what could be the gain brought by PLEDGE when the sources of interest are less or much less powerful than the ones which are known. A hint can be found in Fig.1-(b) which reveals that the θ_2's estimation has been improved. To be sure, one can observe on this figure that when the source of interest is half as powerful as the known one, having then $SNR_2/SNR_1 = 0.5$, the gain is significant. For equipowered sources the advantage could be used at low SNR, less than 10dB, where the gain can nearly reach 5dB. PLEDGE is thus a valuable solution to improved the estimation accuracy into the threshold. We also show that PLEDGE could be much less sensitive to the correlation than MODE and naturally than P-MUSIC which is suboptimal, to be convinced see Fig.2. Lastly, we can see on Fig.1-(c) that the P-MUSIC is better than PLEDGE at low SNR. For this SNR range, P-MUSIC follows a CRB that we named the PLEDGE uncorrelated CRB. That CRB takes into account the prior knowledge of uncorrelated sources and the knowledge of θ_1 (or generally n_k known DOA's), which are the implicit hypothesis of the P-MUSIC algorithm. Due to the lack of space and since this bound is not the key contribution of this work, no closed-form expression of this bound is presented in this paper, but will be left to future work. These additional assumptions on prior-information introduce the next reflexion.

6.2. On prior-knowledge

We develop herein a study on using prior-knowledge. We base our reflexion on the CRBs plotted. Compared to the CRB for n parameters of interest, the PLEDGE CRB which is of reduced dimension, shows that the prior-information is useful especially at low SNR. In addition, increasing the number of known parameters should increase that difference. Pay now attention to the CRB for uncorrelated sources. Make use of this prior hypothesis leads to derive the FIM with $2n + 1$ parameters instead of $n^2 + n + 1$ for the CRB. The number of FIM elements for the PLEDGE is $n_u + n^2 + 1$. Then we could have thought that knowing the sources are uncorrelated would improve more significantly the variance. We can verify this with the help of Fig.1-(c). However for the 2 DOA's case, we can have a scenario from which knowing some directions is better than knowing the uncorrelation state of the sources. That fact is illustrated by Fig.1-(d). We conclude by saying that the minimum variance is obtained when we couple together both the assumption of uncorrelated sources and the assumption of known directions. The corresponding derivation leads to the PLEDGE uncorrelated CRB. This bound has by far the best gain, whatever be the scenario. It is then necessary to give the PLEDGE for uncorrelated sources. This future work will be the optimal prior-knowledge algorithm taking into account both the correlation and the direction. We can imagine furthermore that this new algorithm could be a new tool to test the assumption of correlated/uncorrelated sources.

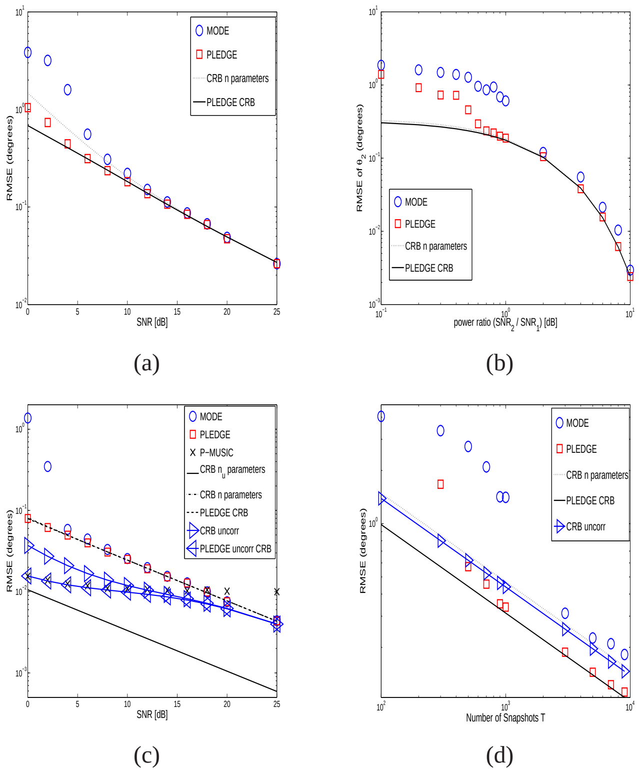

Fig. 1. (a) 5 sensors, 1000 snapshots and equipowered sources, (b) $SNR_1 = 4$ dB, 10 sensors and 100 snapshots, (c) 10 sensors, 100 snapshots $SNR_1 = 4$ dB and $SNR_2 = 14$ dB, (d) SNR=2 dB, 6 sensors and equipowered sources, for all experiments $n = 2$.

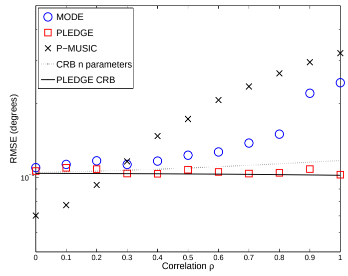

Fig. 2. 10 sensors, 100 snapshots SNR = 2 dB, $\theta_1 = 10°$, $\theta_2 = 14°$ with equipowered sources and $n = 2$.

7. UTILITY FOR MECHANICAL DIAGNOSIS

The mechanical diagnosis seems to be an attractive application by essence. Indeed, many rotating systems generate frequencies linked to the rotation of gears or rotor and stator in induction motors. Thus, acquiring a vibratory signal or a current signal reveals information on the mechanical structure and health of the system. In addition, thanks to the kinematic, some frequencies are known and either do not bring out useful information or prevent the frequencies of interest from being clearly estimated. The application concerned corresponds to the second case mentioned and we will depict the utility of PLEDGE on the estimation of two sideband components masked by a much powerful one. Those frequencies are besides of major importance in rotor bars diagnosis [5].

The following explanations are strongly inspired from http://www.ieee-kc.org/library/motors/motorslip.htm. An induction motor consists of two basic assemblies : a stator and a rotor. The name "induction motor" comes from the alternating current induced into the rotor via the rotating magnetic flux produced in the stator. Motor torque is developed from the interaction of currents flowing in the rotor bars and the stators' rotating magnetic field. In actual operation, rotor speed always lags the magnetic field's speed, allowing the rotor bars to cut magnetic lines of force and produce useful torque. This speed difference is called slip speed. The slip is therefore defined as $s = \frac{v_s - v}{v_s}$ expressed in percent where v is the actual speed and v_s the synchronous speed (stator). For industrial induction motors, the slip is of 1% or 2%. As it is developed in [5], a useful indicator of broken rotor bars is the sideband components around the on line frequency. These sidebands are located in the current spectrum at frequencies $f_s = (1 \pm 2s) f_{line}$ where f_{line} is the network frequency, say the 50Hz. Accordingly, the frequencies of interest are closed from the on line frequency. A drawback of the diagnosis relying on the current signal is the inevitable high dynamic of the 50 Hz. Another drawback is caused by the stability of the network frequency. In France, the electrical producer company ensures the stability of the network frequency on a day (with a certain tolerance only). To estimate the sideband components, the Fourier transform needs a long observation time to increase the precision. However, due to the fluctuation of the 50 Hz, the effect expected is exactly the opposite. The more data you acquire, the less accurate the estimation. That is illustrated by Fig.3. From this figure the sideband frequencies would have been visible on each side of the 50 Hz lobe but due to the instability of the network frequency the Fourier analysis is ineffective. The problem is then to acquire a weak number of data and to have the best precision by possibly getting rid of the 50 Hz which both does not bring out information and is awkward for the estimation. So, PLEDGE seems to be exactly what we need and even if this algorithm has been derived for the array processing we will show that it is suitable in the spectral analysis case.

7.1. Experimental conditions

We have used the current signal acquired from an induction motor (http://www.laspi.fr). The machine as weak broken bars and thus two sideband components should appear at frequencies 51.17 Hz and 48.83 Hz. The experimental conditions are the following : we have acquired 60000 samples, say 100 revolutions, at the sample frequency $f_e = 25.6$ KHz of the current signal and we have decimated of a factor 200 leading to 3 samples per revolution. PLEDGE needs an estimation of the signal subspace by eigendecomposition of the sample covariance matrix. In the context of the spectral line search, we use a Hankel matrix formed by the data samples instead and the MDL technic to estimate the dimension of the signal subspace, see [1] and the references therein. The MDL estimator has given 30 parameters of interest. The 50 Hz network frequency is first estimated by MODE and we use this estimation as the known frequency in PLEDGE.

7.2. Results

The frequencies estimated by PLEDGE are reported inside Table1. The analysis of the table shows that PLEDGE has totally suppressed the known frequency, that is to say the 50 Hz. We focus now on the gray cells which indicate the 2 nearest frequencies from 50 Hz estimated by PLEDGE. Without ambiguity, we can claim that the frequencies 48.87 Hz and 50.98 Hz are those ones which correspond to the slip frequencies. Then, we can surely say that the induction motor has bars defects. So, we have shown that PLEDGE was able to solved the difficult problem of estimating the slip frequencies by getting rid of the 50 Hz and still keeping a good precision. Consequently, from the analysis of the current signal, PLEDGE seems to be adapted for the diagnosis of rotor bars. In addition, few samples are required to monitor the system which is undeniably an advantage to avoid the bad effects caused by the network frequency deviation and to improve the computational cost of the processing.

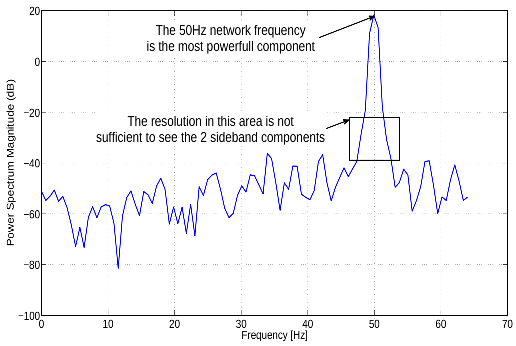

Fig. 3. Power Spectral Density (PSD) of current signal with 300 samples acquired with a sample frequency equals to 128 Hz.

8. CONCLUSION

In this paper we have proposed an optimal **P**rior-know**LEDGE** (**PLEDGE**) method based on the method of direction estimation (MODE) approach. PLEDGE exploits the information of known DOA's to better estimate the unknown ones. To show the benefit of incorporating known information we have also proposed the corresponding stochastic CRB. The simulation results showed that PLEDGE could significantly help the estimation of unknown DOA's especially when the sources corresponding to the unknown DOA's are much less powerful than those ones which are known. We have shown that PLEDGE has good performance into the threshold and is robust to the correlation between the sources, even at low SNR. Finally, we have presented an induction motor diagnosis problem. The diagnosis consists of estimating two slip frequencies masked by a closely located and much more powerful one. We have shown that PLEDGE is very adequate to solve this problem.

Estimated frequencies [Hz] by PLEDGE					
0.15	1.24	6.13	9.76	14.74	15.64
18.05	22.61	23.71	25.17	25.68	26.76
30.04	31.70	34.19	35.49	37.86	39.12
41.20	42.09	45.45	48.87	50.98	52.26
54.69	57.83	58.93	61.92	62.67	63.49

Table 1. The gray cells correspond to the 2 sideband frequencies (slip frequencies)

9. ACKNOWLEDGMENT

G. Bouleux thanks the Rhone-Alpes regional council for having financially supported his stay at the university of Uppsala, Sweden and R. Boyer thanks both the Ile-de-France regional council and Digeteo Research Park.

10. REFERENCES

[1] G. Bouleux, A. Ibrahim, F. Guillet, and R. Boyerb. A Subspace-Based Rejection Method for Detecting Bearing Fault in Asynchronous Motor. In *IEEE International Conference on Condition Monitoring and Diagnosis (CMD 08)*, Beijing, 2008.

[2] R. Boyer and G. Bouleux. Oblique Projections for Direction-Of-Arrival Estimation with Prior Knowledge. *IEEE Trans. on Signal Processing*, 56(4):1374–1387, April 2008.

[3] Y. Bresler and A. Macovski. Exact Maximum Likelihood Parameter Estimation of Superimposed Exponential Signals in Noise. *IEEE ASSP Magazine*, 34:1081–1089, 1986.

[4] H. Chen, S. V. Huffel, D. V. Ormondt, and R. de Beer. Parameter Estimation with Prior Knowledge of Known Signal Poles for the Quantification of NMR Spectroscopy Data in the Time Domain. *Journal of Magnetic Resonance*, 119(2):225–234, 1996.

[5] C. Concari, G. Franceschini, and C. Tassoni. Differential Diagnosis Based on Multivariable Monitoring to Assess Induction Machine Rotor Conditions. *IEEE Trans. on Industrial Electronics*, 55(12):4156–4166, Dec. 2008.

[6] A. B. Gershman and P. Stoica. On Unitary and Forward-Backward MODE. *Digital Signal Processing*, 9:67–75, 1999.

[7] M. Jansson, B. Göransson, and B. Otersten. A Subspace Method for Direction of Arrival Esimation of uncorrelated Emitter Signals. *IEEE Trans. on Signal Processing*, 47(4):945–956, April 1999.

[8] D. Linebarger, R. DeGroat, E. Dowling, P. Stoica, and G. Fudge. Incorporating a Priori Information Into MUSIC-Algorithms and Analysis. *Signal Processing*, 46(1), 1995.

[9] B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai. *Exact and Large Sample ML techniques for Parameter Estimation and Detection in Array Processing*. Radar Array Processing (second ed.). Springer-Verlag, New-York, 1999.

[10] P. Stoica and M. Jansson. On Forward-Backward MODE for Array Signal Processing. *Digital Signal Processing*, 7:239–252, 1997.

[11] P. Stoica, E. G. Larsson, and A. B. Gershman. The Stochastic CRB for Array Processing : A Textbook derivation. *IEEE Signal Processing Letters*, 8(5):148–150, May 2001.

[12] P. Stoica and K. Sharman. Maximum Likelihood Methods for Direction Of Arrival Estimation. *IEEE Trans. on Acoustics Speech and Signal Processing*, 38:1132–1143, 1990.